Interacting with Dynamically Defined Information Spaces using a Handheld Projector and a Pen

Xiang Cao, Ravin Balakrishnan
Department of Computer Science
University of Toronto
caox | ravin @dgp.toronto.edu

ABSTRACT

The recent trend towards miniaturization of projection technology indicates that handheld devices will soon have the ability to project information onto any surface, thus enabling interfaces that are not possible with current handhelds. We explore the design space of dynamically defining and interacting with multiple virtual information spaces embedded in a physical environment using a handheld projector and a passive pen tracked in 3D. We develop techniques for defining and interacting with these spaces, and explore usage scenarios.

ACM Classification: H5.2 [User Interfaces]: Interaction styles. I.3.6 [Methodology and Techniques]: Interaction techniques.

General terms: Design, Human Factors

Keywords: Handheld projector, pen input, information spaces

INTRODUCTION

Handheld devices are increasingly becoming ubiquitous in our daily life: PDAs and cell phones help us manage personal information, access information resources, communicate with others, and entertain ourselves anywhere. However, the small form factor that makes them so appealing is also a significant limitation in that the resulting small screen sizes make it difficult to display large amounts of information, enable rich interactions, or support multiple co-located users viewing the information.

A possible solution to this small screen limitation of handheld devices may lie in recent advances in projection technology, which have seen projectors become smaller, lighter, cheaper, and require less power (Figure 1a). Given this trend, it is reasonable to expect that within a few years projectors could be carried in a pocket or even embedded in other mobile devices such as cell phones and PDAs. With the ability to project information, a handheld device can surmount the limitations of its small internal screen by creating a larger information display on an external surface.

Furthermore, if the projector's position and pose information is available, the system could change the displayed information accordingly, and create an illusion of exploring large information spaces embedded in the physical environment as if using a flashlight (Figure 1b, 2). By doing so, the display and interaction space can be expanded to cover almost an entire physical environment, and support interaction concepts that are not possible on traditional desktop or handheld devices.

Building on these unique affordances of handheld projectors, we explore techniques for dynamically defining and interacting with multiple information spaces embedded in the physical environment using a handheld projector. To enrich the interaction possibilities and leverage human ability to perform bimanual tasks, a passive pen is also used to support annotations and local interactions. We also explore several usage scenarios supported by the system affordances. Although previous research exists on interactive handheld projectors [4, 12, 21, 22], there has not been a systematic exploration of the design space of handheld projector interaction, especially in interactive definition and manipulation of multiple information spaces.

In our prototype we use a commercial camera based motion tracking system to track the projector and the pen with high precision and low latency. However, this is simply an enabler that allows us to explore the interactions of the future with today's technology. Upcoming wireless location tracking systems such as indoor GPS enable such tracking more cheaply and ubiquitously, and their precision and update rate are moving towards what are currently only possible with expensive motion trackers such as the Vicon system we use. For example, Patel et al. [18] and Summet et al. [25] both present handheld devices that can locate themselves in 3D without heavily instrumenting the environment. With ubiquitous location tracking, possibly coupled with embedded orientation sensors in the projectors, we anticipate that our interaction ideas will become practical for wide deployment in the near future.
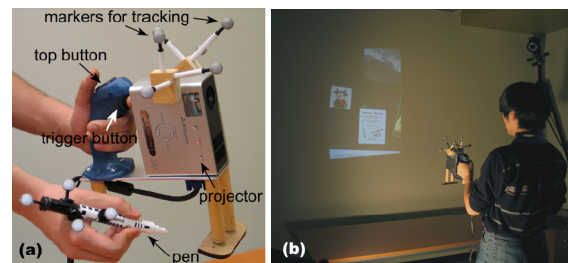
Figure 1. (a) Handheld projector and pen. (b) System in use.

RELATED WORK

To enable interaction with handheld projectors, a key issue is to produce a stabilized projected image that is stationary relative to the projection surface, regardless of the movement of the projector. To achieve this, the projected image is continuously warped to compensate for movement and distortion by taking into account the projector's position and pose. Beardsley et al. [4] and Raskar et al. [22] introduced a computer vision based solution by attaching a camera to the handheld projector to detect visual markers on the projection surface, and corrects the projected image accordingly. This approach requires attaching visual marker to the surface beforehand, thus is not feasible for ubiquitous use in arbitrary environments. Rapp et al. [21] introduced using orientation sensors embedded in the handheld projector to achieve the image correction. This approach does not require installations in the environment, hence can presumably work anywhere. However, the dependence on orientation sensors and lack of knowledge of the environment precludes distinguishing between different projection surfaces, or fixing the displayed information's absolute position in the physical environment. In contrast, our work is the first that allows the user to dynamically define and interact with multiple information spaces embedded in the physical environment.

Researchers [4, 21] have also explored interaction techniques using handheld projection devices. Blaskó et al. [7] investigated embedding a touch-sensitive screen in a simulated wrist-worn projection display to browse stock information. Forlines et al. [12] present Zoom-and-Pick, an interactive widget to address the problem of jittery input and poor image resolution. The Hotaru system [24] supports basic annotation, rotation and file transfer by finger touching on simulated mobile projections. These mainly focus on adapting traditional desktop applications and interaction schemes to handheld projection devices.

The Everywhere Displays project [19, 20] explored using steerable projectors installed in the environment. Peephole Display [27] and Boom Chameleon [26] support browsing and interacting with a large information space using spatially aware displays, similar to the flashlight metaphor in our system. Bimanual interaction has been widely explored [6, 10, 16], including support for multi-user multi-display interaction in an pre-instrumented environment [14]. We explore bimanual use of projector and pen to interact with multiple information spaces without pre-instrumenting the environment. Augmented reality [3, 5, 9, 18] systems superimpose auxiliary information in physical environments. Though conceptually similar to our techniques for authoring information in the environment, the interaction mechanisms and usage scenarios are different. Researchers have also investigated co-located communication between people using sharable displays [17, 23], and public displays that support indirect asynchronous communication [8, 13]. The handheld projector in our system may similarly serve as a semi-public display to support synchronous and asynchronous communication between people.

SYSTEM OVERVIEW

The central concept of our interaction scheme is a flashlight metaphor. The image projected on a physical surface by the handheld projector reveals a portion of a large virtual information space that is situated on the surface (Figure 2a). When the projector is moved, the projected image content changes accordingly to reflect the change of the projected region relative to the information space. By doing so, we create an illusion of exploring a stationary information space relative to the physical surface. When multiple information spaces correspond to different physical surfaces (Figure 2b), this results in a user experience of looking around in a dark room with a flashlight. For the sake of simplicity, from now on we will refer to the virtual information spaces as "spaces", and the corresponding physical projection surface as "surfaces".
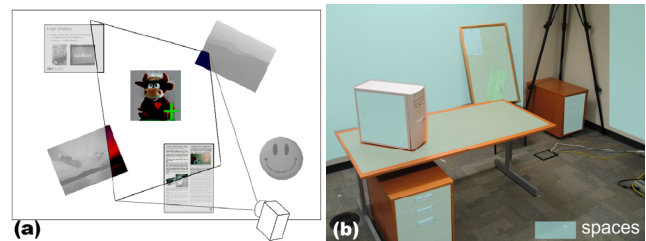

Figure 2. (a) Flashlight metaphor: projected image reveals a portion in the information space. (b) Multiple information spaces in a physical environment.

Virtual objects such as pictures, documents and folders can be contained in the spaces. To indicate the attention point of interaction, a cross-shaped cursor is positioned at the center of the projected image. The cursor can be freely moved across different information spaces by pointing the projector. The cursor size scales according to the distance between the projector and the surface, so as to maintain a relatively constant visual size for the user. Using the cursor and two buttons attached on the projector, the user can interact with the space and the virtual objects. A passive pen is to draw and perform local interactions. The system responds when the pen tip is touching a surface.

Hardware

We use a Mitsubishi™ PK10 Pocket Projector to prototype our system (Figure 1a). The projector weighs about 1 pound, and has a resolution 800×600 pixels. We augmented the projector with a handle and two buttons (trigger button and top button). An integrated stand enables placing the projector on the table in a comfortable projection angle. The passive pen is made from a whiteboard marker with the tip replaced by plastic, and without any electronics embedded. Passive reflective markers attached to the projector and the pen enable tracking by a Vicon camera-based tracking system (www.vicon.com) which provides 6-dof (position + pose) information for both the projector and the pen at millimeter precision at up to 120Hz. The projector is connected to a 2.4GHz Pentium 4 PC, which produces the projection image and handles the interaction.

Algorithms

The projector's optical model is represented by a projection matrix that describes the transformation between a 2D point on the projector's image plane and its projection in the 3D world. An information space is represented as a 3D plane that describes the near-planar surface it corresponds to, and the valid region of the space (represented as a rectangle) that information and interaction must reside within. At each frame, using the projector model, the information space model, and the current position and pose of the projector, a linear system is solved to calculate the projected image. The system then decides what information to project and warps the image to compensate for the distortion caused by non-orthogonal projection (Figure 3).

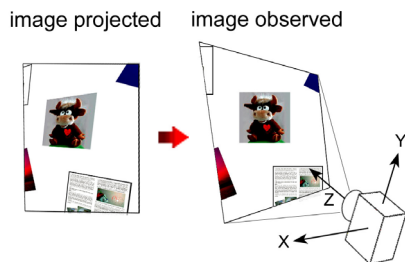image projected   image observed

Figure 3. Display mechanism.

For clarity, we define the local X, Y and Z axes of the projector as in Figure 3. Certain interaction techniques are associated with rotating the projector along some axes.

The projector is calibrated by detecting correspondences between 2D image points and 3D projection points, while the spaces can be defined interactively during use, as described later. To decide whether the pen is being used, the system checks the distance between the pen tip and the surfaces. The pen is active when its tip is less than 3mm from a surface and falls within a valid information space.

A dynamic filter reduces the input jitter caused by hand movement and tracking noise, without noticeable lag.

DESIGN PRINCIPLES

In our exploration, we are guided by the following goals:

Generic interaction schemes: Instead of designing for specific applications, we aim for a set of generic techniques that would be useful for any handheld projector application. Depending on the scenarios, users could combine these low-level interaction techniques to establish higher-level dialogs that are suited to the context.

Applying real-world protocols: Since the interaction spaces are integrated into the physical environment, we adopt protocols that people already use to interact with the physical world and other people where applicable, such as moving towards an object to see details, and using one's body to block other people from seeing private information.

Division and integration between projector and pen: To leverage the different affordances of the projector and the pen, we deliberately assign tasks that are more global and coarse to the projector, and local and precise tasks to the pen. However, where appropriate we also exploit concurrent bimanual interaction with both devices.

Supporting multiple people: Although we focus on the case where only one user operates the projector at a time, the sharability of projected information naturally encourages multiple people to view and work with it. Our designs attempt to accommodate multiple people where possible.

Dynamically defining spaces: In order for the system to be applicable everywhere, the user must not rely on prior knowledge of the environment, but be provided a way to easily and quickly define information spaces on the fly.

DEFINING INFORMATION SPACES

When entering a new environment, the user first defines information spaces corresponding to the physical surfaces in that environment. Any near-planar surface can be potentially defined into a space, e.g. walls, tables, boards, or multiple surfaces of a physical object. To define a space, the system needs to collect two pieces of information: (a) the 3D plane that approximates the surface; (b) the valid region of the space.

Depending on the reachability of the surface and the precision needed for the space, the user may choose to define a space either using the projector or using the pen.

Defining a Space using the Projector

To define a new space, the projector projects a set of concentric quadrangles, each notated with text indicating the desired side length of it, ranging from 1mm to 100m (Figure 4a). These quadrangles stay static relative to the projected image region, just like being projected by an ordinary non-interactive projector. The user's task is to point to the surface to be defined, and adjust the projector pose to make the quadrangles appear as upright squares. At the same time, the user can enlarge the squares by pressing the top button, or shrink them by pressing the trigger button in order to roughly match the displayed sizes of the squares to those suggested by the notations. Alternatively, the user may also move the projector towards or away from the surface to finely adjust the displayed sizes. Once satisfied, the user pressed the two buttons together, and the surface plane of the new space is defined accordingly.

Then the user defines the valid region by sketching the four border lines of the rectangle. Pressing down and holding the trigger button, the user starts to sketch the top border line using the cursor. Releasing the button finishes the line. Similarly, the user then sketches the remaining border lines in the order of right, bottom and left, and the valid region of the space is defined accordingly. The sketched border line segments do not have to form a closed rectangle.
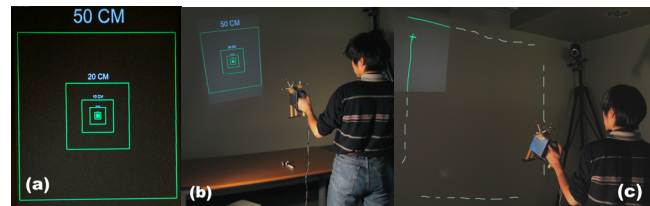

Figure 4. Defining a space using the projector. (a) Projected pattern. (b) Defining the surface. (c) Defining valid region (dotted lines added for illustration purposes only)

We now explain the internal mechanism for this definition procedure. In the surface plane definition stage, the system assumes that the surface plane is always perpendicular to the Z axis of the projector (namely the optical axis) with a fixed distance D from the projector's optical center (Figure 5). Virtual squares are drawn on this imaginary plane with physical sizes indicated by the text notations, and centered at the point that the projector's Z axis intersects the plane. Since the position and orientation of the imaginary plane is updated along with the projector, the projected squares stay static relative to the projection image on the real surface, and distort along with the projection image. By making the displayed quadrangles squares, the user effectively makes the imaginary plane parallel to the real surface to be defined. Pressing the trigger/top button increases/decreases D, i.e. moves the imaginary plane along the projector's Z axis, and results in changes in the sizes of the projected squares on the real surface. By the time the displayed size matches that indicated by the notation, the imaginary plane matches the real surface exactly. Pressing the two buttons together locks the position of the imaginary plane, and defines it to be an information space.
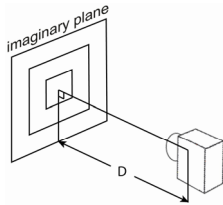

Figure 5. Internal mechanism for defining space using the projector.

After the space is defined, the rectangle of the valid region is calculated by fitting a straight line to each of the border lines the user sketches.

This procedure is applicable to virtually all near-planar surfaces present in the environment, with the exception of the rare cases when the user cannot possibly point the projector perpendicular to it, such as a screen hanging in the air above the user. The accuracy of this procedure is limited by the user's subjective estimation of the squares' shapes and sizes. However, the definition error is negligible if the user only needs to interact with the space using the projector, especially when the surface is far from the user. In order to use the pen on the surface, a more precise definition is required, which can be achieved using the pen itself.

Defining a Space using the Pen
When the surface is within physical reach, the user can define the space using the pen. To do so, the user sketches the four border lines of the space (Figure 6), similar to the region definition stage using the projector. Instead of using the projector cursor, the user sketches the border lines using the pen tip touching the surface. Again, the user uses the trigger button to indicate start and end points of each border line. After border lines are sketched, both the surface plane and the valid region are defined accordingly in a single step.

Since the pen tip is tracked in 3D, the internal mechanism is simply fitting a plane to all the pen tip points collected

during the sketching. The valid region calculation is the same as that in the definition procedure using the projector.
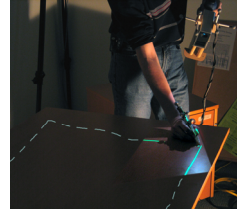

Figure 6. Defining a space using the pen. (dotted lines added for illustration purposes only)

These two definition techniques can be used in combination to efficiently define multiple spaces, and accommodate different surface characteristics. For example, to precisely define a large wall, the user may first define the wall plane using the pen, and then use the projector to define the border lines that are not reachable by the pen. These techniques can also be used to revise an existing space, by simply re-specifying the relevant parts.

INTERACTION TECHNIQUES
In the following sections we describe techniques for interacting with information spaces after they are defined.

Projector-Only Interactions
Using the projector alone, the user can perform tasks that focus on browsing and organizing information.

Virtual Object Manipulation
The basic manipulation of virtual objects is similar to that in a desktop GUI. A click of the trigger button selects the object at the cursor position. Holding the trigger button, the user can move the object around by moving the cursor. Releasing the trigger button releases the object. The object can be seamlessly moved either within the same space, or between different spaces.

In addition, the user can rotate the projector along its Z-axis (effectively by rotating the wrist) to rotate the captured object. Moving and rotating can be performed simultaneously (Figure 7). Since the orientation of the object is consistently mapped to that of the projector, the object stays aligned with the user's view angle, even when it is moved across different spaces, or when the user walks around the space. On the other hand, the user can efficiently position and rotate the object to accommodate other people's perspectives.
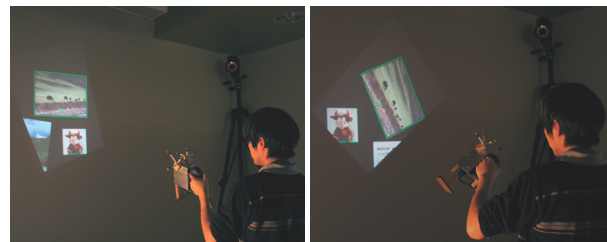

Figure 7. Moving and rotating objects with projector.

Crossing Widgets
Interactive widgets are used to trigger commands and adjust parameters. To provide robustness against the relative lack of

precision of free-hand pointing with the projector, which could be worsened by a physical button-click action, we aim at a set of widgets that alleviate the requirement of precise cursor positioning and button-clicks. We adopted the concept of crossing-based widgets [1, 2] in our projector interaction to achieve facile interaction.

A context-sensitive menu is triggered by pressing the top button. Holding the button, the user uses the cursor to cross a menu item to activate it, and crossing the menu in the opposite direction deactivates the item, as suggested by the green and red edges of the menu (Figure 8a). Releasing the button executes the currently active item, and dismisses the menu. Hierarchical menus are also supported by crossing the items in different layers sequentially.

Crossing based sliders are used to adjust continuous parameters. Crossing the slider activates it, and then the sliding block moves to match the cursor movement parallel to the slider. The parameter value changes accordingly. Note that instead of grabbing the slider block, the user can cross anywhere on the slider to directly dial the block to that position. Crossing the slider in the opposite direction locks the slider block. Using sliders arranged in a row, multiple parameters can be adjusted in a continuous manner. In this case, the previously active slider is locked once another slider is activated. If the user does not want to change some of the parameters, s/he simply moves the cursor around the corresponding sliders to avoid activating them. Figure 8b illustrates using crossing sliders to adjust the size and transparency of an object. Incidentally, crossing sliders are also used as scrollbars for scrollable objects such as documents or presentation slides.
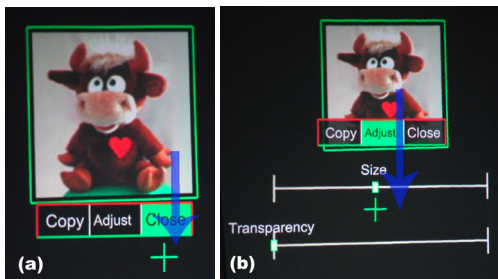

Figure 8. Crossing widgets. (a) Crossing menu. (b) Crossing sliders (arrows show the crossing path).

By associating different sliders (or groups of sliders) with different menu items, we provide a fluid way to adjust many parameters in a single continuous sequence of crossing actions without need for numerous button-clicks.

These interactive widgets automatically move and scale, so that they always stay inside the projected image. Hence the user can operate the widgets using local movements, while reposition them within/between spaces using more global movements, without need for a mode switch. Tracking menus [11] provide similar affordances to operate and reposition a widget with a single pointer. However, since our design takes into account the projected image region, the result is a more implicit and less intrusive mechanism. In addition, it smoothly accommodates the user's movement in the physical environment. Because the size of the projected image is proportional to the distance between the user (i.e. the projector) and the surface, scaling the widgets accordingly results in approximately the same visual size and the same amount of hand movement to operate the widgets.

When invoked, the widget also adjusts its orientation (0°, 90°, 180° or 270°) to accommodate the user's view angle, according to the orientation of the projector at that moment.

Resolution Gradation and Information Granularities
A unique characteristic of using a handheld projector is that the local image resolution changes depending on the distance between the projector and the surface. When the projector moves closer to the surface, the same amount of image pixels is distributed in a smaller region, resulting in a higher local resolution. This naturally matches user's experience of viewing physical spaces. From afar the user has a large but low-resolution display region to overview the information space, while coming closer, the user focuses on a small but high-resolution region to acquire details. This resolution gradation spontaneously happens in an implicit continuous manner.

In addition to image resolution gradation, dynamic resolution can be applied to the semantics of information. A virtual object may switch between different information granularities when the distance between the user and the surface changes. High-level information is displayed when the user is far from the surface, and detailed information is displayed when the user comes closer. This is realized by calculating the area ratio between the virtual object itself and the projected image region, which effectively measures how many image pixels are distributed to the object, given the total amount of pixels of the projector is fixed. This also results in a switch of information granularities while scaling the object. Figure 9 illustrates a city map switching between granularities when the user moves.

Figure 9. Proximity adjusts information granularity.

Space Overview
Since the projector usually displays only a portion of the information space at one time, the user may occasionally have difficulties in navigating the space and locating a virtual

object. To assist navigation, pressing both buttons together triggers an overview panel showing a miniature of the current space (Figure 10). Like the interactive widgets, the panel moves and scales with the projected image region. When moved into a different space, the content of the panel switches to overview the newly entered space.

Figure 10. Space overview panel.

Flick Gesture
Depending on the context, a "flick" gesture to the left or the right acts as a shortcut to frequently used menu commands. To make the flick gesture, the user quickly rotates the projector along its Y-axis to the left or right and then back again (Figure 11a). To inform the user about the shortcut, the corresponding menu item is marked with a gesture icon (Figure 11b). Figure 11c illustrates using a flick gesture to page up/down in a document, and Figure 11d illustrates using it to switch between functions (magnifying, increasing contrast, and querying information) for a multi-functional magic lens. Incidentally, in the information query mode, the information text rotates along with the magic lens, providing a way to adjust the text orientation to accommodate different view angles, either for different people, or for the same user at different times.
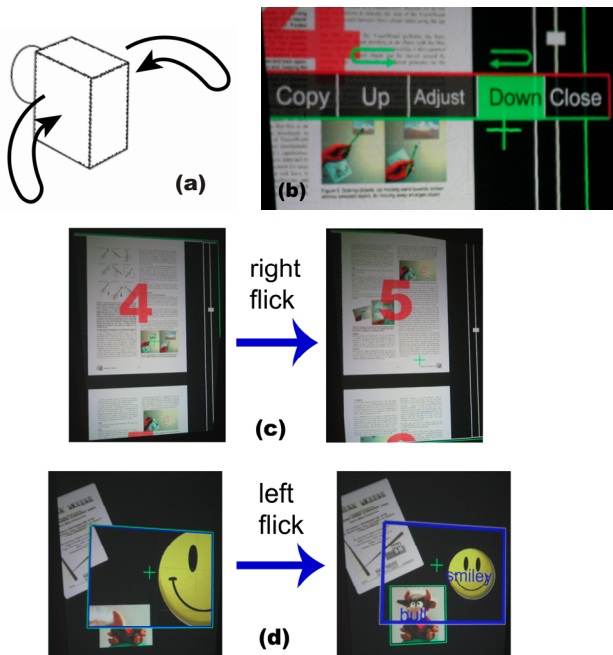

Figure 11. (a) Flick gesture. (b) Gesture icons in a menu. (c) Turning pages. (d) Switching lenses.

Personal Folder
A personal folder contains the virtual objects that the user may want to interact with. Through a menu command, this personal folder can be accessed in any space, and the objects inside it can be dragged into the spaces.

Depending on the usage scenario, these virtual objects might be stored in the handheld device (personal data), in devices in the environment (context data), or both.

Pen and Projector Interactions
Using the passive pen along with the handheld projector, the user can draw pen inks in information spaces, annotate on virtual or physical objects, and perform local interactions more precisely.

Although most of the pen techniques described below do not necessarily require holding the projector (the projector could be put down on a table, or not used at all in an "eyes-free" scenario), we suspect users will hold the projector with the non-dominant hand to set the display/interaction context, and use the dominant hand to perform pen interactions. Therefore these pen-based techniques are described as bimanual interaction techniques.

When the pen tip touches the surface, a pen stroke can be drawn in the space or on a virtual object (Figure 12). Neighboring strokes are grouped into ink. Ink can be moved, rotated, scaled (using a crossing-slider) and closed (using a crossing-menu) just like other virtual objects. The pen does not have to reside inside the projected image region to draw, hence the user may make "blind" notes while viewing other portion of the spaces, or make "secret notes" when s/he does not want other people to see what is being written.
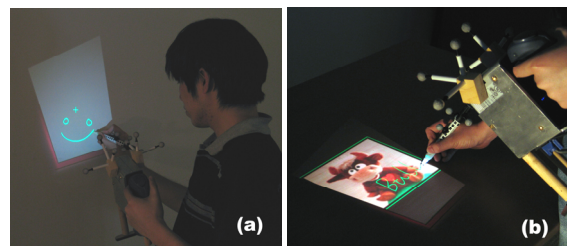

Figure 12. (a) Drawing on a surface. (b) Annotating a virtual object.

In order to annotate on a remote virtual object that is out of physical reach, the user can use the projector cursor to capture and drag the object to a closer location while holding the trigger button, and then annotate on it using the pen. To avoid jittery inking caused by the movement of the hand holding the projector, the object's position is frozen once the user starts writing. Incidentally, when the object is too large to be displayed completely by the projector, this enables the user to temporarily pin down the object with the pen and move the projector to browse it. Once the user finishes the annotation, s/he releases the trigger button, and the annotated object flies back to its original position, as if it is spring-loaded. This provides an efficient way to annotate virtual objects scattered around a large physical environment, without the need to walk around or rearrange the objects.

Utilizing the functionality of drawing in spaces, the user can also annotate static physical objects in the environment. Depending on the relationship between the physical object and the information space, the object may be annotated in

three different ways: (a) on the surface of the object, when the surface has been defined into an information space. This applies to relatively large objects with near-planar surfaces, such as a box or a computer. (b) Over the object, when it forms part of an information space. This applies to flat objects attached to other surfaces, such as a poster or a brand. (c) Around the object, when it protrudes from an information space. This applies to relatively small objects that stay on top of other surfaces, such as a toy on the table, or a switch on the wall. Figure 13 illustrates.
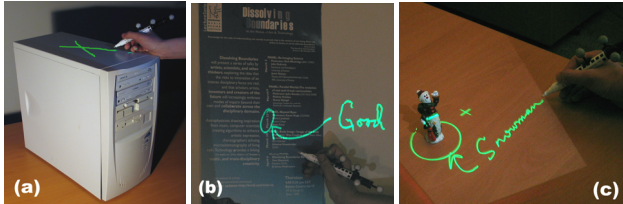

Figure 13. Annotating physical objects.

The pen can also be used to perform local interactions with a virtual object, which enables precise control that is difficult to achieve with the projector cursor. Figure 14 illustrates using the pen to operate the crossing-based scrollbar of a document. Note that the pen does not have to act inside the projected image. Once the user activates the scrollbar, s/he can focus the projector on the document and scroll it using the pen, without paying visual attention to the widget any more. Complete eyes-free operation of the widget is also possible if the user has become sufficiently familiar with its position. This enables the user to focus on the information content rather than the interface.

Figure 14. Using the pen to scroll a document.

Although we describe these pen-based techniques as bimanual interactions, they may also be used for multi-user interaction where one user controls the projector to set the context, and another uses the pen to perform the actions.

USAGE SCENARIOS

Utilizing the interaction techniques described, various user scenarios can be supported. Instead of focusing on specific applications, we categorize these scenarios by the generic interaction patterns, both between the user and the system, and between different people.

Single-Person Use

The capability to easily and quickly define information spaces anywhere enables the user to temporarily convert the physical environment around into spaces to manage personal information (Figure 15). Depending on the environment, these spaces can take various forms, such as walls, tables, floors, seatbacks, or even the user's laps. The dynamic image

size and resolution of the handheld projector readily accommodates interacting on surfaces of different sizes and distances. The possible large scale of the spaces supports tasks that are difficult on a traditional handheld device with limited display and interaction space, such as managing a personal album, browsing a map, or annotating a large document. When the user leaves the environment, s/he closes the information spaces, and takes away the processed information along with the device.

Spatial memory can help the user to efficiently organize information, especially when multiple information spaces are available. For example, in an office room the user may place different groups of information on different walls, and designate the ceiling as a recycle bin. By doing so, the user can easily locate the information, and use coarse hand movement (possibly eyes-free) to organize the information between walls. In this particular case, moving a virtual object upward to the ceiling performs "send to recycle bin", and effectively acts as a gesture-like command.
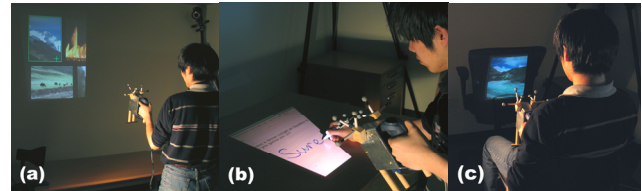

Figure 15. Single user scenarios. (a) Managing photo album. (b) Annotating a document. (c) Interacting on the back of a seat.

Synchronous Multi-Person Use

Our techniques can also support communication between multiple people. We focus on the case where one main user controls the projector and shows the information, while other people watch the information and use the pen to communicate. Scenarios that involve use of multiple projectors together will be explored in the future.

Instead of supporting formal meetings in a dedicated room, the handheld projector is more suited for casual ad hoc collaborations. Similar to the single-user scenario, people can temporarily annex their physical environment into information spaces to support their communication. Even a simple task like showing information to others can utilize different interactions depending on context (Figure 16). Note that privacy concerns may arise when the main user needs to process personal information that s/he does not wish other people to see. In this case, moving close to the surface results in a small display region, which can be blocked from others using the user's own body.

Multiple people can also use the pen to facilitate communications. A blank wall can be turned into a virtual whiteboard to support brainstorming (Figure 17a). People can also collaboratively annotate a virtual object (e.g. editing a document) or a physical object. Figure 17b illustrates two people discussing about a physical poster. When the main user is far from others (e.g. a teacher in a classroom), s/he can use the projector to pass on a virtual writing pad (or any other virtual objects) to them and let them write on it (Figure

17c). Once the main user releases the button, the writing pad flies back to him/her with the notes made by other people. This provides an efficient way to collect comments from people in a large environment.
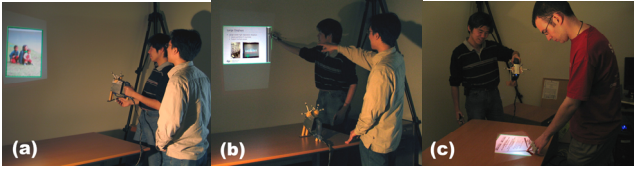

Figure 16. Synchronous multi-user scenario. (a) Main user holds the projector, showing a picture for both people. (b) Showing presentation slides, with the projector on the table and using the pen to switch slides. (c) Main user positions and orients the document to accommodate the other person.
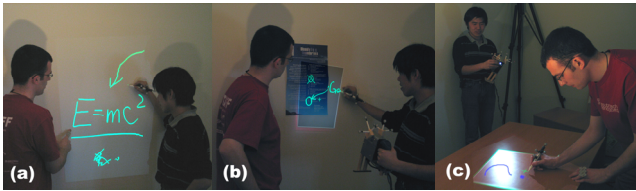

Figure 17. Projector and pen used in multi-user scenario. (a) Virtual white board. (b) Discussing a poster. (c) Passing a writing pad.

Asynchronous Multi-Person Use

Instead of closing the information spaces when people leave, the information may also be left in the physical environment to facilitate asynchronous communication. When another user enters the environment, s/he can use the projector to retrieve information left by others. Thus, the handheld projector serves as an information conduit between people across time.

Provided the ability to define any near-planar surface into an information space, the user can leave pen inks and paste information anywhere in the physical environment. When another user comes into the environment later, s/he can use the projector to retrieve the information left in the environment by the other people (Figure 18). The relationships between physical objects in the environment and the virtual information can be two-way. The virtual information may serve as annotations and auxiliary information for a physical object (Figure 18c), or the physical object may serve as a proxy to attract attention to the virtual information (Figure 18a, b).

The concept of overlaying information in physical environments has been explored in augmented reality applications [3, 5, 9]. However, our ability to easily define spaces and author the overlaying information on the fly may result in casual asynchronous communication between users. Multiple people may also create a rich information environment through asynchronous collaboration, similar to collaborative information sharing on public displays [8, 13]. This scenario also includes the case that a single user leaves reminders to him/herself for future use.
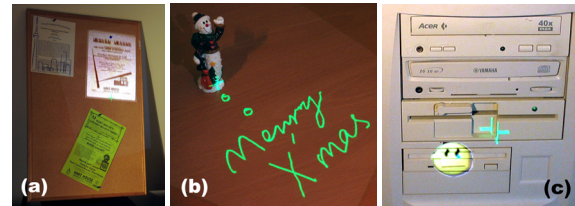

Figure 18. Information left in the environment. (a) Virtual flyer on a physical bulletin board. (b) Note written on a desk. (c) Error icon on broken device.

Game Use

The affordances of the system also open up design options for games. Figure 19 illustrates a mockup of a shooting game. With the guidance of stereo sound effects suggesting the positions of virtual enemies, the player uses the projector to spot and shoot at them. Other game scenarios are possible, such as using the pen as a gestural input device to cast spells, as if using a magic wand. The player can dynamically define the physical game environment, and exploit the physical layout of the environment to diversify the game experience. The result is an alternate game setting that can be deployed virtually anywhere.

Figure 19. Shooting game.

INFORMAL USER FEEDBACK

In order to collect some early informal user feedback, we asked five graduate students to try the prototype system. Each participant was shown the interaction and space definition techniques, and asked to freely and extensively try out the techniques and functionalities. Each trial session lasted 40-50 minutes. Only the single user scenario was tested. We observed participants' behaviors, and informally interviewed them about their opinions and suggestions, and in what scenarios they would use the system.

All participants easily understood the flashlight metaphor and the presence of multiple information spaces. Their first impressions of the system were unanimously "very cool". They did not show or express any difficulties in learning the interaction techniques. The overview widget was found to be especially helpful. Though one participant raised the concern of possible fatigue resulting from moving the projector all the time, nobody actually reported fatigue during the trial periods. Although no participants had prior experience with crossing interfaces, the crossing-based widgets were broadly welcomed. They felt that these widgets not only greatly reduced the problem of jittery input, but also saved time in performing complex operations. All participants were enthusiastic with the ability to leave virtual annotations in the physical environment. They were particularly excited about

the idea of leaving "secret" notes, since this reminded them of "magic holographs" in fairy tales.

The main complaints relate to the technical limitations of the system, such as display jitter, having to manually adjust the projector's focus, and the limited working area of the tracking system. Another concern is that because only part of the information space is displayed, sometimes the user may lose the interaction context, such as when browsing a menu with deep hierarchy, or performing a flick gesture. One possible solution is to let the context temporarily move with the projector in certain situations.

The participants made suggestions for improvements. Two participants expressed the desire to incorporate handwriting recognition to input text. Some participants suggested adding more input channels on the projector, such as a small keyboard or joystick, as present on cell phones. One user asked for the ability to write in the air using the pen.

The participants also conceived many scenarios in which they would use the system, e.g. taking notes, browsing and editing large documents, giving presentations, collaboratively working with others, looking up for context-related instructions, and entertainment.

DISCUSSION

Although we have not yet conducted an iterative user centered (re)design, we carefully considered the particular characteristics of our system when making design choices. For example, we believe that two physical buttons on the projector provide sufficient input bandwidth for the interface, and suspect that additional buttons could add unnecessary complexity from both physical ergonomics and ease of comprehension perspectives. The adoption of crossing interfaces instead of point-and-click interfaces was motivated by the observations that freehand pointing in 3D space is somewhat imprecise.

Although seeing only a portion of the information spaces seems to be a limitation in some cases, in other senses it may also act as an advantage. The ability to choose what and where to display provides the user control over the information spaces. Therefore we can avoid the information overload problem well-known in ubiquitous computing [15], when information is all over the place and overwhelms the user's attention. Furthermore, when multiple people are present, this also avoids disturbing other people with irrelevant information, and supports privacy for personal information.

In the scenario of asynchronous communication between people, the user needs to discover the information posted in the environment previously. This may be supported in various ways. When the user is seeking information about a physical object, s/he would actively browse around it with the projector. On the other hand, physical objects can serve as proxies to attract people's attention for information update, e.g. a bulletin board with both physical and virtual posters on it, or a small object deliberately left on the desk. Automatic reminders provided by the system may also help, e.g. when the user enters the environment, visual arrows or audio notifications could guide him/her to check for information updates.

If a projector can be uniquely identified by the system, e.g. with embedded RFID, user-specific information can be accessed without needing to log in. On one hand, this supports more directed asynchronous communication between people. The user could leave information that is addressed to a specific person. On the other hand, different layers of information can be accessed by different people at the same time in a public environment.

Our system provides some affordances similar to augmented reality (AR) systems (e.g., [18]), such as annotating physical objects. However, the use of a handheld projector makes our system more appropriate for collaborative applications, compared with AR devices which are designed typically for single-person use. Our system also eliminates the need for prior knowledge of the environment, which is essential for similar systems using situated projectors such as Everywhere Displays [19, 20].

CONCLUSIONS AND FUTURE WORK

This work has explored interaction scenarios for a handheld projector and a pen. The ability to easily define and interact with multiple virtual information spaces embedded in physical environments enables the system to be used ubiquitously to support personal information management as well as communication between people. Informal user feedback indicates the interaction design is reasonable and useful. Our work is an attempt to define generic interaction vocabularies and user scenarios for practical use of interactive handheld projection devices, and is also the first to apply pen-based and bimanual interaction along with a handheld projector.

Although we have investigated multi-user scenarios using one projector at a time, it would certainly be interesting to explore the case where each user has a handheld projector, and the interaction between those projectors at the same time. We also plan to support information spaces embedded in moving surfaces (such as a handheld notepad) and non-planar surfaces (such as human bodies). By attaching a camera to the handheld projector, it could not only place information into the physical environment, but also capture information from the environment. When situated large displays are available in the environment, the handheld projector could be used to overlay information on them, resulting in multi-layer information spaces. On the theoretical side, we are also interested in investigating people's ability to navigate across multiple spaces embedded in a 3D environment with partial information visible at one time.

REFERENCES

1. Accot, J. and Zhai, S. (2002). More than dotting the i's - foundations for crossing-based interfaces. *ACM CHI Conference on Human Factors in Computing Systems*. p. 73-80.

2. Apitz, G. and Guimbretière, F. (2004). CrossY: a crossing-based drawing application. *ACM UIST Symposium on User Interface Software and Technology*. p. 3-12.

3. Barfield, W. and Caudell, T., eds. *Fundamentals of wearable computers and augmented reality*. 2001, Lawrence Erlbaum Associates: Mahwah, New Jersey.

4. Beardsley, P., Baar, J.V., Raskar, R., and Forlines, C. (2005). Interaction using a handheld projector. *IEEE Computer Graphics and Applications*, 25(1). p. 39-43.

5. Bell, B., Feiner, S., and Höllerer, T. (2001). View management for virtual and augmented reality. *ACM UIST Symposium on User Interface Software and Technology*. p. 101-110.

6. Bier, E., Stone, M., Pier, K., Buxton, W., and DeRose, T. (1993). Toolglass and Magic Lenses: The see-through interface. *ACM SIGGRAPH Conference on Computer Graphics and Interactive Techniques*. p. 73-80.

7. Blaskó, G., Coriand, F., and Feiner, S. (2005). Exploring interaction with a simulated wrist-worn projection display. *ISWC'05 IEEE International Symposium on Wearable Computers*. p. 2-9.

8. Brignull, H., Izadi, S., Fitzpatrick, G., Rogers, Y., and Rodden, T. (2004). The introduction of a shared interactive surface into a communal space. *ACM CSCW Conference on Computer Supported Cooperative Work*. p. 49-58.

9. Buxton, W., *Living in augmented reality: Ubiquitous Media and Reactive Environments*, in *Video Mediated Communication*, K. Finn, Sellen, A., Wilber, S., Editor. 1997, Erlbaum: Hillsdale, NJ. p. 363-384.

10. Buxton, W. and Myers, B. (1986). A study in two-handed input. *ACM CHI 1986 Conference on Human Factors in Computing Systems*. p. 321-326.

11. Fitzmaurice, G., Khan, A., Pieke, R., Buxton, B., and Kurtenbach, G. (2003). Tracking menus. *ACM UIST Symposium on User Interface Software and Technology*. p. 71-79.

12. Forlines, C., Balakrishnan, R., Beardsley, P., Baar, J.v., and Raskar, R. (2005). Zoom-and-pick: facilitating visual zooming and precision pointing with interactive handheld projectors. *ACM UIST Symposium on User Interface Software and Technology*. p. 73-82.

13. Greenberg, S. and Rounding, M. (2001). The notification collage: posting information to public and personal displays. *ACM CHI Conference on Human Factors in Computing Systems*. p. 514-521.

14. Hilliges, O. and Terrenghi, L. (2006). Overcoming mode-changes on multi-user large displays with bi-manual interaction. *MU3I Workshop on Multi-User and Ubiquitous User Interfaces*.

15. Intille, S.S. (2002). Change blind information display for ubiquitous computing environments. *Ubicomp*. p. 91–106.

16. Matsushita, N., Ayatsuka, Y., and Rekimoto, J. (2000). Dual touch: A two-handed interface for pen-based PDAs. *ACM UIST Symposium on User Interface Software and Technology*. p. 211-212.

17. Morris, M.R., Ryall, K., Shen, C., Forlines, C., and Vernier, F. (2004). Beyond "social protocols": multi-user coordination policies for co-located groupware. *ACM CSCW Conference on Computer Supported Cooperative Work*. p. 262-265.

18. Patel, S., Rekimoto, J., and Abowd, G. (2006 - in press). iCam: Precise at-a-distance interaction in the physical environment. *Pervasive*.

19. Pingali, G., Pinhanez, C., Levas, A., Kjeldsen, R., Podlaseck, M., Chen, H., and Sukaviriya, N. (2003). Steerable Interfaces for Pervasive Computing Spaces. *PerCom'03 IEEE International Conference on Pervasive Computing and Communications*. p. 315--322.

20. Pinhanez, C. (2001). The everywhere display projector: A device to create ubiquitous graphical interfaces. *Ubicomp*. p. 315--331.

21. Rapp, S., Michelitsch, G., Osen, M., Williams, J., Barbisch, M., Bohan, R., Valsan, Z., and Emele, M. (2004). Spotlight navigation: Interaction with a handheld projection device. *International Conference on Pervasive Computing, Video paper*.

22. Raskar, R., Beardsley, P., Baar, J.v., Wang, Y., Dietz, P., Lee, J., Leigh, D., and Willwacher, T. (2004). RFIG lamps: interacting with a self-describing world via photosensing wireless tags and projectors. *ACM Transactions on Graphics*, 23(3). p. 406-415.

23. Simon, A. (2005). First-person experience and usability of co-located interaction in a projection-based virtual environment. *ACM Symposium on Virtual reality software and technology*. p. 23-30.

24. Sugimoto, M., Miyahara, K., Inoue, H., and Tsunesada, Y. (2005). Hotaru: Intuitive Manipulation Techniques for Projected Displays of Mobile Devices. *INTERACT 2005*. p. 57-68.

25. Summet, J. and Suthankar, R. (2005). Tracking locations of moving hand-held displays using projected light. *Pervasive*. p. 37-46.

26. Tsang, M., Fitzmaurice, G.W., Kurtenbach, G., Khan, A., and Buxton, B. (2002). Boom chameleon: simultaneous capture of 3D viewpoint, voice and gesture annotations on a spatially-aware display. *ACM UIST Symposium on User Interface Software and Technology*. p. 111-120.

27. Yee, K.-P. (2003). Peephole displays: pen interaction on spatially aware handheld computers. *ACM CHI Conference on Human Factors in Computing Systems*. p. 1-8.